PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 5/10, 15/09, 15/12, 15/62, 15/63, 15/79, C07K 19/00, 14/705	A1	(11) International Publication Number: WO 95/34641 (43) International Publication Date: 21 December 1995 (21.12.95)

(21) International Application Number: PCT/US95/07542

(22) International Filing Date: 13 June 1995 (13.06.95)

(30) Priority Data:
08/260,514 15 June 1994 (15.06.94) US

(71) Applicant: THE SCRIPPS RESEARCH INSTITUTE [US/US]; 10666 North Torrey Pines Road, TPC-8, La Jolla, CA 92037 (US).

(72) Inventors: MUIR, Thomas, W.; 3745 Ingraham Street, San Diego, CA 92109 (US). WILLIAMS, Michael, J.; 4029 Carmel View Road, #149, San Diego, CA 92130 (US). GINSBERG, Mark, H.; 2944 Renault Place, San Diego, CA 92122 (US). KENT, Stephen, B.H.; 6046 La Jolla Hermosa Avenue, La Jolla, CA 92037 (US). CHEN, Yi-Ping; 8155 Cargill Avenue, #40, San Diego, CA 92122 (US). O'TOOLE, Timothy, E.; 4600 Lamont #4-326, San Diego, CA 92109 (US).

(74) Agents: LOGAN, April, C. et al.; The Scripps Research Institute, 10666 North Torrey Pines Road, TPC 8, La Jolla, CA 92037 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: STRUCTURAL MODELS FOR CYTOPLASMIC DOMAINS OF TRANSMEMBRANE RECEPTORS

(57) Abstract

The present invention comprises model systems that allow study and control of the activity of the intracellular or cytoplasmic regions of integrins and other transmembrane proteins. One aspect of these model systems is protein constructs that contain helical regions in order to approximate the effect of the transmembrane hydrophobic regions while maintaining solubility. One type of protein construct according to the present invention comprises: (1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and (2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment. The protein construct has one free alpha-amino terminus derived from one of the amino acid sequences of the first segment and two free carboxyl termini. Another type of protein construct according to the present invention comprises a first segment, two second segments, and two third segments. Also within the scope of the invention are chimeric integrins and methods for their use.

Serial No.: 09/144,838
Inventor: Siani et al.
Attny Docket No. GRFN-020/01US

F6

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgystan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SI	Slovenia
CI	Côte d'Ivoire	KZ	Kazakhstan	SK	Slovakia
CM	Cameroon	LI	Liechtenstein	SN	Senegal
CN	China	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
CZ	Czech Republic	LV	Latvia	TJ	Tajikistan
DE	Germany	MC	Monaco	TT	Trinidad and Tobago
DK	Denmark	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	US	United States of America
FI	Finland	ML	Mali	UZ	Uzbekistan
FR	France	MN	Mongolia	VN	Viet Nam
GA	Gabon				

STRUCTURAL MODELS FOR CYTOPLASMIC DOMAINS OF TRANSMEMBRANE RECEPTORS

BACKGROUND OF THE INVENTION

This invention is directed to structural models for cytoplasmic domains of transmembrane receptors, including protein constructs and chimeric proteins, as well as methods of synthesis and use of such structural models.

In eukaryotic cells, many proteins extend through the membrane and therefore have a cytoplasmic domain, a transmembrane domain, and an extracellular domain. Many of these proteins are involved in signal transduction, the control of cell adhesion, and cell-cell interaction control.

Among the proteins that fall into this category are the integrins (R.O. Hynes, "Integrins: Versatility, Modulation, and Signalling in Cell Adhesion," Cell 69:11-25 (1992); S.M. Albelda & C.A. Buck, "Integrins and Other Cell Adhesion Molecules," FASEB J. 4: 2868-2880 (1990); E. Ruoslahti, "Integrins," J. Clin. Invest. 87: 1-5 (1991); M.E. Hemler, "VLA Proteins in the Integrin Family: Structures, Functions, and Their Role on Leukocytes," Annu. Rev. Immunol. 8:365-400 (1990)).

Integrins are involved in a number of pathological and physiological processes, including thrombosis, inflammation, and cancer (M.J. Humphries et al., "A Synthetic Peptide From Fibronectin Inhibits Experimental Metastasis of Murine Melanoma Cells," Science 233:467-470 (1986); E. Ruoslahti, *supra*: 1-M.E. Hemler (1990), *supra*). Other physiological and pathological conditions involving changes in cell adhesiveness are also mediated through integrins.

Typically, transmembrane proteins are heterodimeric, being noncovalent associations of two or more different types of polypeptide subunits. In particular, integrins are heterodimers of two different protein subunits, designated α and β. The α subunits vary in size between 120 and 180 kDa and are each noncovalently associated

with a β subunit. The extracellular domain of the integrin molecule forms a ligand binding site; both the α and β subunits are involved in forming the ligand binding site. A number of different ligands for integrins are known are known, including collagens, laminin, fibronectin, vitronectin, complement components, thrombospondin, and integral membrane proteins of the immunoglobulin superfamily such as ICAM-1, ICAM-2, and VCAM-1. The integrins recognize various short peptide sequences in the ligands. Examples of these are Arg-Gly-Asp (RGD), Lys-Gln-Ala-Gly-Asp-Val (KQAGDV) (SEQ ID NO: 1), Asp-Gly-Glu-Ala (DGEA) (SEQ ID NO: 2), Glu-Ile-Leu-Asp-Val (EILDV) (SEQ ID NO: 3), and Gly-Pro-Arg-Pro (GPRP) (SEQ ID NO: 4). Variations in integrin function are often caused by changes in the ligand binding affinity of the extracellular domain of the integrins (J.S. Bennett & G. Vilaire, J. Clin. Invest. 64: 1393-1401 (1979); D.C. Altieri et al., J. Cell Biol. 107: 1893-1900 (1988); R.J. Faull et al., J. Cell Biol. 121: 155-162 (1993); B.A. Lollo et al., J. Biol. Chem. 268:21693-21700 (1993)).

Integrin $\alpha_{IIb}\beta_3$ (platelet GPIIb-IIIa), a heterodimer of two type I transmembrane protein subunits (R.O. Hynes (1992), *supra*), manifests highly regulated changes in ligand binding affinity. Affinity state-specific antibodies, e.g., PAC1 (S.J. Shattil et al., J. Biol. Chem. 260:1107-1114 (1985)), are useful for analysis of recombinant $\alpha_{IIb}\beta_3$ in heterologous cells (T.E. O'Toole et al., Cell Regulation 1:883-893 (1990)). Platelet agonists increase the affinity of $\alpha_{IIb}\beta_3$ (activation) probably by causing changes in the confirmation of the extracellular domain (T.E. O'Toole et al. (1990), *supra*; P.J. Sims et al., J. Biol. Chem. 266:7345-7352 (1991)). Cytoplasmic signaling pathways involving heterotrimeric GTP binding proteins, phospholipid metabolism, and serine-threonine kinases initiate these confirmational changes in the extracellular domain; these changes may also involve calcium fluxes, tyrosine kinases, and low

molecular weight GTP binding proteins (P.J. Sims et al. (1991), *supra*; S.J. Shattil et al., J. Biol. Chem. 267: 18424-18431 (1992); S.J. Shattil & J.S. Brugge, Curr. Opin. Cell Biol. 3: 869-879 (1991); M.H. Ginsberg et al., Cold Spring Harbor Symposium of Quantitative Biology: The Cell Surface 57: 221-231 (1992); M.H. Ginsberg et al., "Inside-Out Integrin Signalling," Curr. Opin. Cell Biol. 4: 766-771 (1992); Y. Nemoto et al., J. Biol. Chem. 267: 20916-20920 (1992)). How cytoplasmic signals result in changes in the conformation and ligand binding affinity of the extracellular domain ("inside-out signal transduction") of the integrin remains obscure. Studies with chimeras containing the cytoplasmic domains of various α and β subunits joined to the transmembrane and extracellular domain of $\alpha_{IIb}\beta_3$ indicates that integrin cytoplasmic domains transduce cell type-specific signals that modulate ligand binding affinity. These signals require active cellular processes in both α and β cytoplasmic tails of the integrin, suggesting that they reflect physiologically relevant signals. In addition, deletion of a highly conserved motif, Gly-Phe-Phe-Lys-Arg (GFFKR) (SEQ ID NO: 5), at the amino-terminus of the α subunit cytoplasmic domain, also resulted in high affinity binding of ligands to integrin $\alpha_{II}\beta_3$. In contrast to the chimeras, high affinity ligand binding to GFFKR (SEQ ID NO: 5) deletion mutants was independent of cellular metabolism, cell type, and the bulk of the β subunit cytoplasmic domain. Thus, integrin cytoplasmic tails are targets for the modulation of integrin affinity.

Although signal transduction by integrins, as well as by transmembrane proteins generally, is of great interest, technical difficulties have greatly limited the application of high resolution techniques for the structures of these proteins. Indeed, molecular structures are available for only two intact transmembrane proteins, a bacterial photoreaction center (J. Deisenhofer et al., Nature 318: 618-624 (1985)), and

a porin (M.S. Weiss et al., FEBS Lett. 267: 268-272 (1990)). Structures of receptor extracellular domains have been determined using soluble truncated extracellular domains as models (A.M. DeVos et al., Science 255: 306-312 (1992); M.V. Milburn et al., Science 254: 1342-1347 (1991)). These structures have contributed to the understanding of the basis of ligand recognition, but have provided less insight into the mechanism of signal transduction. Many membrane proteins that transduce signals are members of the Type I transmembrane protein family, the defining feature of which is a single membrane spanning region. These include the T cell receptor (A. Weiss, Cell 73: 209-212 (1993)); growth factor receptors (L. Patthy, Cell 61: 13-14 (1990), and cytokine receptors (A. Miyajima et al., TIBS 17: 378-382 (1992)). In general, the cytoplasmic domain of these proteins is critical for signaling. Thus, to understand signal transduction through such receptors, it is essential to understand the structure and function of the cytoplasmic domain. This is especially difficult for multisubunit Type I proteins.

An additional complication in studying and in developing methods of controlling or modulating the activity of integrins and other transmembrane proteins is the inherent hydrophobicity of membrane-spanning stretches of apolar amino acids. This means that structural studies on truncated receptors containing these helical regions will be impeded by insolubility in aqueous solution.

Therefore, there is a need for improved methods of controlling and modulating the activity of integrins and other transmembrane proteins, of detecting their activity, and modulating their activity to detect and control physiological conditions.

SUMMARY

To meet these needs, we have developed model systems that allow study and control of the activity of the intracellular or cytoplasmic regions of integrins and

other transmembrane proteins. Particularly, these model systems maintain the cytoplasmic regions in their natural conformation. These model systems include protein constructs and chimeric integrins.

One aspect of the present invention is protein constructs intended to link cytoplasmic regions of transmembrane proteins in a framework providing solubility in aqueous media. In general, a protein construct according to the present invention comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment; the protein construct having: (i) either no free α-amino terminus or one free α-amino terminus derived from one of the two amino acid sequences of the first segment and (ii) two free carboxyl termini.

One protein construct according to the present invention comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment.

The protein construct has one free α-amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini. This is referred to as a Type I protein construct.

Typically, the substantially helical amphiphilic acid regions of the first segment have a predominantly

periodic secondary structure. Preferably, the substantially helical amphiphilic amino acid regions in the first segments have an estimated helicity of at least about 80%; more preferably, the helicity is at least about 85%.

Typically, the substantially helical amphiphilic amino acid regions each have the sequence G-(X_1-L-X_2-X_3-L-X_4-G)$_n$, wherein X_1 is selected from the group consisting of lysine, arginine, and ornithine, X_2 and X_4 are each independently selected from the group consisting of aspartic acid and glutamic acid, X_3 is selected from the group consisting of alanine, serine, and threonine, and n is an integer from 2 to 20. Preferably, n is from 3 to 6; more preferably, n is 4. Preferably, X_1 is lysine, X_2 and X_4 are each glutamic acid, and X_3 is alanine.

Typically, n is identical for both of the substantially helical amino acid regions; preferably, the sequence of both of the substantially helical amphiphilic amino acid regions is identical. However, these are not absolute requirements, and constructs in which n is different for the substantially helical amino acid regions or in which the sequence of the substantially helical amino acid regions differs are within the scope of the invention.

Typically, each of the two second segments has a length of about 15 amino acids to about 50 amino acids. The two second segments can be non-helical. Typically, the amino acid sequences of each of the two second segments are derived from the cytoplasmic domain of a transmembrane protein. More typically, the amino acid sequences of each of the two second segments are derived from the amino acid sequences of the cytoplasmic domains of the subunits of a heterodimeric multisubunit transmembrane protein where the subunits noncovalently associate *in vivo*. The amino acid sequences of each of the two second segments can be derived from the cytoplasmic domain of an integrin. In this case,

preferably the amino acid sequence of one of the second segments is derived from the cytoplasmic domain of an integrin selected from the group consisting of α_1, α_2, α_3, α_4, α_5, α_6, α_7, α_8, α_{IIb}, α_V, α_L, α_M, α_X, and α_{IEL}, and the amino acid sequence of the second of the second segments is derived from the cytoplasmic domain of an integrin selected from the group consisting of β_1, β_2, β_3, β_4, β_5, β_6, β_7, and β_8, so that one of the following combinations is formed: $\alpha_1\beta_1$, $\alpha_2\beta_1$, $\alpha_3\beta_1$, $\alpha_4\beta_1$, $\alpha_4\beta_7$, $\alpha_5\beta_1$, $\alpha_6\beta_1$, $\alpha_6\beta_4$, $\alpha_7\beta_1$, $\alpha_8\beta_1$, $\alpha_V\beta_1$, $\alpha_V\beta_3$, $\alpha_V\beta_5$, $\alpha_V\beta_6$, $\alpha_V\beta_8$, $\alpha_L\beta_2$, $\alpha_M\beta_2$, $\alpha_X\beta_2$, $\alpha_{IIb}\beta_3$, and $\alpha_{IEL}\beta_7$.

A particularly preferred embodiment is one in which the amino acid sequence of one of the second segments is derived from the cytoplasmic domain of integrin α_{IIb} and the amino acid sequence of the second of the second segments is derived from the cytoplasmic domain of integrin β_3. In this embodiment, the amino acid sequence of the first of the second segments can be residues 989-1007 of integrin α_{IIb}, with an additional carboxyl-terminal glutamine, and the amino acid sequence of the second of the second segments can be that of residues 716-762 of integrin β_3.

Preferably, the helicity of at least one of the first and second segments is increased in the construct over the helicity of the first or second segment alone.

Typically, the two amino acid sequences of the first segment are linked by a thioether linkage between a cysteine residue at the amino terminus of one of the amino acid sequences of the first segment and a bromoacetyl moiety at the amino terminus of the second of the amino acid sequences of the first segment. Other linkages can also be used.

In an alternative Type I protein construct according to the present invention, the two amino acid sequences are linked by an oxime linkage between an aldehyde moiety and an aminooxy moiety. This alternative Type I protein construct can comprise:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion through an oxime linkage formed by reaction of an aldehyde moiety with an aminooxy moiety, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment.

This alternative Type I protein construct has either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

One of the second segments can have an amino acid sequence derived from the cytoplasmic domain of α subunit of an integrin with a deletion of a sequence G-F-F-K-R.

A particularly preferred protein construct comprises:

(1) a first segment including two copies of a substantially helical amphiphilic amino acid sequence that is G-(K-L-E-A-L-E-G)$_4$, joined by a thioether linkage formed by reaction of a cysteine residue that is linked to the amino-terminal end of one of the two amino acid sequences and a bromoacetyl moiety that is linked to the amino-terminal end of the second amino acid sequence; and

(2) two second segments, each covalently linked to a carboxyl terminus of the first segment, wherein one of the second segments has the amino acid sequence of residues 989-1007 of integrin α_{IIb} with an additional carboxyl-terminal glutamine residue, and the second second segment is residues 716-762 of integrin β_3.

Another embodiment of the invention is a protein construct referred to herein as a Type II protein construct, which comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion so that the first segment has no free amino terminus and two carboxyl termini available for covalent linkage to the amino-terminal end of another amino acid segment;

(2) two second segments, each including a substantially helical amphiphilic amino acid region, each of the second segments being covalently linked at their amino-terminal ends to a carboxyl terminus of the first segment; and

(3) two third segments, each of the third segments being an amino acid sequence, each of the third segments being covalently linked by their amino-termini to the carboxyl-termini of the second segment, the protein construct having no free amino terminus and two free carboxyl termini derived from the third segments.

The second and third segments of this Type II protein construct are analogous to the first and second segments of the Type I protein construct, but the Type II protein construct has additional residues between the helical regions and the covalent linkage. The first segment can include a specific binding partner sequence having affinity for a specific binding partner. The specific binding partner sequence can specifically bind an antibody.

Another aspect of the present invention is a method of producing protein constructs. For Type I protein constructs using a thioether linkage, this method can comprise the steps of:

(1) providing two amino acid sequences, each amino acid sequence including: (i) an amino-terminal amino acid segment; (ii) an intermediately situated substantially helical amphiphilic amino acid segment covalently linked at its amino terminus to the first segment; and (iii) a

third amino acid segment covalently linked at its amino terminus to the second amino acid segment within the amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through a thioether linkage to form a protein construct having one free α-amino terminus derived from one of the two amino acid sequences and two free carboxyl termini.

For Type I protein constructs using an alternative oxime linkage, the method can comprise the steps of:

(1) providing two amino acid sequences, each amino acid sequence comprising a substantially helical amphiphilic amino acid region at its amino terminus covalently linked to a second amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through an oxime linkage between an amino acid residue with an aldehyde moiety and an amino acid residue with an aminooxy moiety to form a protein construct having either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

For Type II protein constructs using a thioether linkage, the method can comprise the steps of:

(1) providing two amino acid sequences, each amino acid sequence including: (i) an amino-terminal amino acid segment; (ii) an intermediately situated substantially helical amphiphilic amino acid segment covalently linked at its amino terminus to the first segment; and (iii) a third amino acid segment covalently linked at its amino terminus to the second amino acid segment within the amino acid sequence; and

(2) joining the two amino acid sequences in head-to-

head fashion through an oxime linkage between an amino acid residue with an aldehyde moiety and an amino acid residue with an aminooxy moiety to form a protein construct having either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

For alternative Type II protein constructs using an oxime linkage, the method can comprise the steps of:

(1) providing two amino acid sequences, each amino acid sequence including: (i) an amino-terminal amino acid segment; (ii) an intermediately situated substantially helical amphiphilic amino acid segment covalently linked at its amino terminus to the first segment; and (iii) a third amino acid segment covalently linked at its amino terminus to the second amino acid segment within the amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through an oxime linkage between an amino acid residue with an aldehyde moiety and an amino acid residue with an aminooxy moiety to form a protein construct having either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

Another aspect of the present invention is chimeric

integrin proteins and methods for their use.

Chimeric integrin proteins according to the present invention can comprise: (1) a chimeric integrin protein comprising the extracellular and transmembrane domains of the Tac subunit of the IL-2 receptor covalently linked to the cytoplasmic domain of integrin β_3; (2) a chimeric integrin protein comprising the extracellular and transmembrane domains of the Tac subunit of the IL-2 receptor covalently linked to the cytoplasmic domain of integrin β_3 with amino acid 752 being mutated from a serine residue to a proline residue; (3) a chimeric integrin protein comprising the transmembrane and extracellular domains of the Tac subunit of the human IL-2 receptor covalently linked to the cytoplasmic domain of integrin α_{IIb}; and (4) a heterodimeric chimeric integrin in which the extracellular and transmembrane domains of human integrin $\alpha_{IIb}\beta_3$ are joined to the cytoplasmic domains of human integrin $\alpha_5\beta_1$.

Another aspect of the invention is nucleic acid sequences encoding the chimeric integrin proteins. Yet another aspect of the invention is these nucleic acid sequences operatively linked to at least one control element for transcription of the nucleic acid sequence, as well as vectors comprising the nucleic acid sequences operatively linked to the control elements, the vectors being capable of transfecting at least one eukaryotic host for expression of the chimeric integrin proteins encoded by the nucleic acid sequences.

Another aspect of the invention is a method for blocking the activation of a human cellular integrin comprising the step of expressing a chimeric integrin whose extracellular and transmembrane domains are derived from the Tac subunit of the human IL-2 receptor and whose cytoplasmic domain is the cytoplasmic domain of the human β_1 integrin in a quantity sufficient to inhibit high affinity ligand binding by the cellular integrin. Alternatively, in this method, the chimeric integrin

protein comprising the extracellular and transmembrane domains of the Tac subunit of the IL-2 receptor covalently linked to the cytoplasmic domain of integrin β_3 can be used.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features, aspects, and advantages of the present invention will become better understood with reference to the following description, appended claims, and the accompanying drawings where:

Figure 1A is a schematic drawing of the presumed topology of the native integrin $\alpha_{IIb}\beta_3$;

Figure 1B is a schematic drawing of a Type I protein construct of the present invention designated MP-1;

Figure 2 is a depiction of the amino acid sequences of the synthetic polypeptide targets, including a helix-dimer, MP-1, and MP-2, showing the relationships between the sequences in the assembled molecules;

Figure 3 is a schematic equation showing the synthesis of MP-1 using chemical dovetailing;

Figure 4 shows graphs of analytical HPLC results illustrating the effect of reaction conditions on formation of the thioether bond in the chemical dovetailing step as a function of time: (4(a), aqueous solution (0.1 M phosphate, pH 7); 4(b), dimethylformamide-0.1 M phosphate (95:5), pH 7);

Figure 5 shows the results of studies intended to characterize the assembled polypeptides MP-1 (5(a)) and MP-2 (5(b)); main panel, ion-spray mass spectra showing charge state distribution; insets, analytical reverse phase HPLC spectra;

Figure 6 shows the results of circular dichroism studies of synthetic polypeptides containing the individual cytoplasmic domains of the integrin $\alpha_{IIb}\beta_3$ (6(a), isolated cytoplasmic domains, 6(b), cytoplasmic domains attached to helical segments);

Figure 7 shows the results of circular dichroism studies of MP-1, MP-2, and the covalent helix-dimer;

Figure 8 shows the results of fluorescence quenching

studies on MP-1, MP-2, isolated β_3, and isolated β_3 + isolated α_{IIb};

Figure 9 is a schematic depiction of various chimeric integrins and Tac chimeras, including the amino acid sequences of the cytoplasmic regions of the integrins;

Figure 10 shows histograms depicting the results of flow cytometry in which fluorescence intensity is depicted on the abscissa and cell number on the ordinate; in Panels A, B, and C, PAC1 binding in the absence (filled histogram) or presence (open histogram) of competitive inhibitor is shown; in Panel D, surface expression of $\alpha_{IIb}\beta_3$ is shown as reported by the binding of D57 antibody in the absence (filled histogram) or presence (open histogram) of co-transfected Tac-β_3 DNA; M1 indicates the region containing those cells that express the recombinant $\alpha_{IIb}\beta_3$ construct bearing the cytoplasmic domains of β_1 and α_5;

Figure 11A is a graph showing the inhibition of inside-out signaling by Tac chimeras ((■) β_3, (●) β_1, (▲) α_5) used to transfect CHO cells, with signaling assessed by PAC1 binding;

Figure 11B is a graph showing the expression of the Tac chimeras used in Figure 11A assessed by the binding of the anti-TAC antibody 7G7B6;

Figure 12A is a graph showing the effect of the mutation $S^{752}P$ on inside-out signaling in cells transfected with chimeras as in Figure 11A ((▲) Tac-β_3($S^{752}P$); (■) Tac-β_3);

Figure 12B is a graph showing the expression of the Tac chimeras used in Figure 11A assessed as in the experiments shown in Figure 11A and 11B;

Figure 13A is a graph showing the effect of a Tac-β_1 chimera on the affinity of a "hinge" mutant, $\alpha_L\Delta$ ((●) α_5; (■) $\alpha_L\Delta$), by cotransfection as in Figure 11A; and

Figure 13B is a graph showing the expression of the Tac-β_1 chimera in the experiments of Figure 13A assessed

as in the experiments shown in Figure 11A and 11B.

DESCRIPTION

Definitions

In the context of this disclosure, including the description, examples, and claims, the following terms are defined as follows unless otherwise indicated:

"Protein Construct" means a molecule constructed from two or more protein segments that are covalently joined in such a way that the molecule has one free α-amino terminus derived from one of the protein segments but two or more free carboxyl termini.

"Nucleic Acid" includes both DNA and RNA unless otherwise indicated, and can include both single and double-stranded nucleic acid sequences. If a DNA sequence is referred to, references generally to both strands of a DNA sequence, either individually or as a Watson-Crick double helix. If only one strand is specified, the complementary strand whose antiparallel sequence is determined by Watson-Crick based pairing rules is also included unless the complementary sequence is specifically excluded. If only one strand is specified in double-stranded DNA, the strand specified as the sense strand whose strand would be equivalent to the sequence of any RNA transcribed from the double-stranded DNA, except for the replacement of thymidine (T) in the DNA by uridine (U) in the RNA. Reference to a nucleic acid sequence also includes modified bases as long as the modification does not significantly interfere with Watson-Crick base pairing or other specified functions of the nucleic acid, and can, for example, include substitution of uridine for thymidine in DNA as well as methylation of bases or modification of sugars.

"Helical" refers to any periodic regular secondary structure of amino acids in a protein, domain of a protein, or peptide characterized by a uniform translation along an axis, and includes but is not limited to an α-helix unless so specified.

"Periodic" refers to any regular secondary structure

of amino acids in a protein that is not a random coil or a quasi-random coil, and includes, but is not limited to, a helical structure.

"Specific binding partner" refers to a member of a pair of molecules that interact by means of specific non-covalent interactions that depend on the three-dimensional structures of the molecules involved, such as salt links, hydrophobic interactions, and van der Waals interactions. Typical pairs of specific binding partners include antigen-antibody, hapten-antibody, hormone-receptor, nucleic acid strand-complementary nucleic acid strand, substrate-enzyme, inhibitor-enzyme, carbohydrate-lectin, biotin-avidin, and virus-cellular receptor.

"Antibody" includes both intact antibodies and antibody fragments possessing an antigen-binding site, as well as chimeric antibodies.

"Amino terminus" refers to an amino group or moiety that is part of an amino acid residue within a protein, polypeptide, or protein construct. The amino group can either be an α-amino terminus or an amino terminus from a side chain, such as the ϵ-amino group of lysine.

"Amino-terminal residue" refers to an entire amino acid residue possessing an amino terminus. The amino-terminal residue can be joined to other amino acid residues, either through its carboxyl terminus by a peptide bond, or through a side chain.

I. PROTEIN CONSTRUCTS

One aspect of the present invention is protein constructs that maintain the cytoplasmic tails of transmembrane proteins in conformations approaching their naturally occurring conformations while preserving solubility in an aqueous medium and obviating the difficulty of working with proteins containing highly hydrophobic transmembrane segments.

These protein constructs allow study of the conformation and activity of the cytoplasmic regions of transmembrane proteins, including integrins, and can be used, as discussed below, in screening and in controlling

the activity of integrins and other transmembrane proteins.

A. Structure of Protein Constructs

In general, a protein construct according to the present invention comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment;

the protein construct having: (i) either no free α-amino terminus or one free α-amino terminus derived from one of the two amino acid sequences of the first segment and (ii) two free carboxyl termini.

1. Protein Construct With First and Second Segments

One aspect of the present invention is a protein construct with first and second segments. In general, this protein construct comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion, each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment. The protein construct has one free α-amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini. For convenience, this protein construct is referred to herein as a Type I protein construct.

The free α-amino terminus derived from one of the two amino acid sequences of the first segment is free in

the sense that it is not involved in a peptide bond. However, unlike the situation in naturally-occurring proteins, the amino acid residue to which this α-amino terminus belongs is covalently linked to two other amino acid residues, although one of the linkages is through a side chain; the other is through the carboxyl terminus of that residue, through a peptide bond. This α-amino terminus derived from one of the amino acid sequences of the first segment must be distinguished from any amino terminus derived from the side chains of the amino acid sequence, typically ε-amino termini derived from lysine side chains.

Typically, the substantially helical amphiphilic acid regions of the first segment have a predominantly periodic secondary structure. More typically, the substantially helical amphiphilic amino acid regions in the first segment have an estimated helicity of at least about 80%. Preferably, the substantially helical amphiphilic amino acid regions in the first segment have an estimated helicity of at least about 85%. One suitable measurement technique for helicity is circular dichroism (CD), but other measurement techniques such as optical rotatory dispersion (ORD) can be used.

Typically, the substantially helical amphiphilic amino acid regions have the sequence is $G-(X_1-L-X_2-X_3-L-X_4-G)_n$. In this repeating sequence, n is an integer from 2 to 20, preferably from 3 to 6, more preferably 4. In the repeating sequence, X_1 is lysine, arginine or ornithine, X_2 and X_4 can each be glutamic acid or aspartic acid, and X_3 can be alanine, serine, or threonine. Typically, the sequence of the two substantially helical amphiphilic amino acid regions is the same, but this is not required. Typically, n is the same for both of the two substantially helical amphiphilic amino acid regions, but this is also not required. In some applications, it may be desirable to provide protein constructs in which the substantially helical amphiphilic regions are of different lengths.

Preferably, X_1 is lysine, X_2 and X_4 are each glutamic acid, and X_3 is alanine. Derivatives of these amino acids that do not alter the relative hydrophobicity or hydrophilicity or disrupt the helix are also included within the scope of the invention.

In one particularly preferred alternative of the invention, the helical amphiphilic amino acid regions form coiled-coil tertiary and quaternary structures (O.D. Monera et al., "Comparison of Antiparallel and Parallel Two-Stranded α-Helical Coiled-Coils," J. Biol. Chem. 268: 19218-19227 (1993); N.E. Zhou et al., J. Biol. Chem. 267: 2664-2670 (1992)) analogous to those of the prototypical coiled-coil protein tropomyosin. The stability of these coiled-coils is largely a result of strong interchain hydrophobic interactions between leucine residues in the seven residue repeat (S.Y.M. Lau et al., J. Biol. Chem. 259: 23253-13261 (1984)). These coiled-coil structures are likely to better mimic the proximity of transmembrane helices in the natural system and also ensure that a defined topology is maintained between the α and β cytoplasmic tails. In other words, the coiled-coil can act as a structural template onto which the cytoplasmic domain of the integrin or other transmembrane protein is attached. This ensures that the two cytoplasmic tails are staggered with respect to one another in a manner that approximates the intact protein.

Typically, each of the two second segments has a length of about 10 amino acids to about 80 amino acids; more typically, each of the two second segments has a length of about 15 amino acids to about 50 amino acids. The two second segments can be predominantly non-helical in structure. The amino acid sequences of each of the two second segments can be derived from the cytoplasmic domain of a transmembrane protein. Typically, the amino acid sequences of the two second segments are derived from the amino acid sequences of the cytoplasmic domains of the subunits of a heterodimeric multisubunit transmembrane protein where the subunits noncovalently

associate *in vivo*, but it is also possible to prepare constructs according to the present invention in which the amino acid sequences of the second segments are derived from protein subunits that do not normally associate into a heterodimeric transmembrane protein *in vitro*.

In one embodiment, the amino acid segments of each of the two second segments are derived from the cytoplasmic domain of an integrin. The protein constructs of the present invention are particularly useful for mimicking integrin conformation because, in a heterodimeric integrin, the location of the transmembrane-cytoplasmic interface of both families of subunits is formed by conserved Trp-Lys or Tyr-Lys residues. These residues are proximal to each other at the interface, so that subunits of differing lengths have a predictable relative stagger. The cytoplasmic tails are connected to helical transmembrane stretches, which may be important for inducing structure along the tails, and both cytoplasmic tails have carboxyl termini.

The following combinations of sequences from the α and β subunits of integrins can be used according to Hynes (1992), *supra*, incorporated herein in its entirety by this reference: Subunit β_1 can associate with any of α_1, α_2, α_3, α_4, α_5, α_6, α_7, α_8, and α_V. The amino acid sequence from β_2 can be associated with the amino acid sequence from α_L, α_M and α_X. The amino acid sequence from β_3 can be associated with α_{IIb} or α_V. The amino acid sequence of subunit β_4 can be associated with α_6. For subunits β_5, β_6, and β_8, the amino acid sequence can be associated with the amino acid sequence of α_V. For β_7, the amino acid sequence can be associated with the amino acid sequence for α_4 or α_{IEL}. The term "associated" in this context means that a heterodimeric integrin in which the cytoplasmic domains of the sequences are as specified is likely to form *in vivo*.

In one particularly preferred embodiment, the amino

acid sequence of one of the second segments is derived from the cytoplasmic domain of integrin α_{IIb} and the amino acid sequence of the other of the second segments is derived from the cytoplasmic domain of β_3.

Preferably, the amino acid sequence of the first of the second segments is derived from residues 989-1008 of α_{IIb} and the amino acid sequence of the second of the second segments is that of residue 716-762 of β_3. The residue 1008, which is glutamic acid (E) in the natural integrin according to the DNA sequence, is preferably altered to glutamine (Q) in the protein construct, because it is believed that post-translational side-chain amidation of the carboxyl-terminal glutamate occurs (J.J. Calvete et al., FEBS Lett. 263:43-46 (1990)).

Although one particularly useful protein construct derives its second segments from sequences of integrins, the sequences of other transmembrane proteins can be used. These include the T-cell receptor, cytokine receptor, other CD antigens, or the growth hormone receptor.

The two amino acid sequences of the first segment are joined at their amino-terminal residues in head-to-head fashion so that there is one free α-amino terminus, as described above. Preferably, the first and second segments are joined through a thioether linkage. Other joining methods also are known in the art.

In an alternative Type I protein construct according to the present invention, the first and second segments are joined through an oxime linkage formed between an aminooxy moiety on the amino-terminal residue of one of the two amino acid sequences of the first segment and an aldehyde residue on the amino-terminal residue of the other amino acid sequence of the first segment. This protein construct comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion through an oxime linkage formed by reaction of an aldehyde moiety with an aminooxy moiety,

each of the two amino acid sequences of the first segment including a substantially helical amphiphilic amino acid region; and

(2) two second segments, each of the second segments joined to the first segment at the carboxyl terminus of each of the two amino acid sequences of the first segment.

The protein construct has either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

In this alternative Type I protein construct, it is not necessary that the linkage of the two amino acid sequences of the first segment be through the side chain of a cysteine residue. If a free α-amino terminus exists in this alternative Type I protein construct, it is available for reaction.

Typically, in the constructs of the present invention, the helicity of at least one of the first and second segments is increased in the construct over the helicity of the first or second segment alone, as measured by circular dichroism.

In another embodiment of the present invention, one of the second segments has an amino acid sequence derived from the α subunit of an integrin, such as α_{IIb}, with a deletion of the sequence GFFKR (SEQ ID NO: 5). This deletion causes an increase of the affinity of the integrin for its extracellular ligand and is sometimes referred to as a "hinge mutation." A schematic of a typical Type I protein construct of the present invention, using second segments whose amino acid sequences are derived from α_{IIb} and β_3 integrins, is shown in Figure 1.

Although the protein constructs disclosed above contain two cytoplasmic domains of transmembrane proteins, related techniques of chemoselective ligation can be used to generate higher-order protein constructs containing three, four, or more cytoplasmic domains of transmembrane proteins held in a conformation approximating the conformation *in vivo*. This technique is known generally as template-assisted protein synthesis, and is described in P.E. Dawson & S.B.H. Kent, "Convenient Total Synthesis of a 4-Helix TASP Molecule by Chemoselective Ligation," J. Am. Chem. Soc. 115: 7263-7266 (1993).

2. Protein Construct With First, Second, and Third Segments

Another embodiment of the present invention is a protein construct comprising:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion so that the first segment has one free α-amino terminus derived from one of the two amino acid sequences of the first segment and two carboxyl termini available for covalent linkage to the amino-terminal end of another amino acid segment;

(2) two second segments, each including a substantially helical amphiphilic amino acid region, each of the second segments being covalently linked at their amino-terminal ends to a carboxyl terminus of the first segment; and

(3) two third segments, each of the third segments being an amino acid sequence, each of the third segments being covalently linked by their amino-termini to the carboxyl-termini of the second segment. The protein construct has one free amino terminus derived from the first segment and two free carboxyl termini derived from the third segments. For convenience, this construct is referred to as a Type II protein construct herein.

The second and third segments of the Type II protein construct are substantially equivalent to the first and

second segments of the Type I protein construct described above. However, the first segment includes additional amino acid residues.

The first segment of the Type II construct can include a specific binding partner sequence having affinity for a specific binding partner. The specific binding partner sequence can specifically bind an antibody.

An alternative Type II protein construct uses an oxime linkage between two amino acid sequences of the first segment as discussed above. This protein construct comprises:

(1) a first segment including two amino acid sequences joined at their amino-terminal residues in head-to-head fashion through an oxime linkage formed by reaction of an aldehyde moiety with an aminooxy moiety so that the first segment has two carboxyl termini available for covalent linkage to the amino-terminal end of another amino acid segment;

(2) two second segments, each including a substantially helical amphiphilic amino acid region, each of the second segments being covalently linked at their amino-terminal ends to a carboxyl terminus of the first segment; and

(3) two third segments, each of the third segments being an amino acid sequence, each of the third segments being covalently linked by their amino-termini to the carboxyl-termini of the second segment.

This alternative Type II protein construct has either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

B. Synthesis of Protein Constructs

Typically, synthesis of a type I protein construct using a thioether linkage according to the present invention occurs by the following process:

(1) providing two amino acid sequences, each amino acid sequence comprising a substantially helical amphiphilic amino acid region at its amino terminus covalently linked to a second amino acid sequence; and

(2) covalently linking two amino acid sequences in head-to-head fashion through a thioether linkage to produce a protein construct having one free α-amino terminus derived from one of the two amino acid sequences and two free carboxyl termini.

Typically, synthesis of an alternative Type I protein construct using an aminooxy linkage occurs by the following process:

(1) providing two amino acid sequences, each amino acid sequence comprising a substantially helical amphiphilic amino acid region at its amino terminus covalently linked to a second amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through an oxime linkage between an amino acid residue with an aldehyde moiety and an amino acid residue with an aminooxy moiety to form a protein construct having either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

Typically, synthesis of a Type II protein construct using a thioether linkage according to the present invention occurs by the following process:

(1) providing two amino acid sequences, each amino acid sequence including: (i) an amino-terminal amino acid

segment; (ii) an intermediately situated substantially helical amphiphilic amino acid segment covalently linked at its amino terminus to the first segment; and (iii) a third amino acid segment covalently linked at its amino terminus to the second amino acid segment within the amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through a thioether linkage to form a protein construct having one free α-amino terminus derived from one of the two amino acid sequences and two free carboxyl termini.

Typically, synthesis of an alternative Type II protein construct using an aminooxy linkage occurs by the following process:

(1) providing two amino acid sequences, each amino acid sequence including: (i) an amino-terminal amino acid segment; (ii) an intermediately situated substantially helical amphiphilic amino acid segment covalently linked at its amino terminus to the first segment; and (iii) a third amino acid segment covalently linked at its amino terminus to the second amino acid segment within the amino acid sequence; and

(2) joining the two amino acid sequences in head-to-head fashion through an oxime linkage between an amino acid residue with an aldehyde moiety and an amino acid residue with an aminooxy moiety to form a protein construct having either zero or one free amino terminus derived from one of the two amino acid sequences of the first segment and two free carboxyl termini, with the proviso that when the protein construct has one free amino terminus, the amino acid residue having the aminooxy moiety has an amino group derived from a side chain, the free amino terminus of the protein construct being derived either from the α-amino terminus or from the amino terminus of the side chain.

The amino acid sequences comprising the first and second segments or the first, second, and third segments are generally synthesized through peptide synthesis. The

principles of peptide synthesis, particularly solid-phase peptide synthesis, are well known in the art and are described, for example, in M. Bodanszky, "Principles of Peptide Synthesis" (2d ed., Springer, Berlin, 1993). Typically, synthesis is performed on a solid phase, and involves activation and coupling, the reversible blocking of amino and carboxyl groups, and eventual deprotection. A particularly useful method of solid phase protein synthesis is described in M. Schnölzer, "In Situ Neutralization in Boc-Chemistry Solid Phase Peptide Synthesis," Int. J. Peptide Protein Res. 40:180-193 (1992), incorporated herein by this reference. This method uses a t-butyloxycarbonyl protecting group.

Alternatively, the amino acid sequences needed for assembly of the constructs can be produced through recombinant expression techniques that yield the desired products. These techniques are well known in the art and need not be described further here.

The synthesized segments are then covalently linked in head-to-head fashion. Preferably, a free cysteine residue on one of the amino acid sequences is condensed with a bromoacetyl moiety on the other amino acid sequence to yield a protein construct with a thioether linkage. Such techniques are described, for example in, T.W. Muir & S.B.H. Kent, "The Chemical Synthesis of Proteins," Curr. Opin. Biotechnology 4:420-427 (1993) and in M. Schnölzer & S.B.H. Kent, Science 256:221-225 (1992), both of which are incorporated herein by this reference.

C. Applications

The protein constructs of the present invention have a number of applications based on the ability to maintain the cytoplasmic tails of the construct in a configuration that is equivalent or similar to the configuration predominating *in vivo* while maintaining solubility and stability in an aqueous system, namely in staggered, parallel, and proximal topology. For example, these protein constructs can be used to detect intracellular

molecules capable of binding to integrins and modulating signals by inside-out signaling. Alternatively, these molecules can be used *in vivo* to disrupt or modulate inside-out signaling by binding to the cells in a manner such that the cytoplasmic domains of these protein constructs compete for intracellular molecules with the natural integrins. Because these protein constructs do not contain the extracellular ligand-binding sites of integrins, they would then disrupt inside-out signaling. This would be particularly useful in conditions in which overactivity of integrins is involved, such as inflammation, thrombosis, and malignancy.

Additionally, protein constructs according to the present invention could be used to detect molecules capable of binding to the intracellular or cytoplasmic domain of integrins and other transmembrane molecules *in vivo*. The protein constructs could be modified with aldehyde or aminooxy groups and cross-linked to affinity columns for chromatography. Affinity chromatography is well known in the art and need not be described here; typical procedures are described, for example in G.T. Hermanson et al., "Immobilized Affinity Ligand Techniques" (Academic Press, San Diego, 1992), incorporated herein by this reference. Other derivatized constructs can be made with various derivatives for affinity chromatography, according to crosslinking reactions well known in the art. Constructs according to the present invention can also be used in screening for oncogens.

II. CHIMERIC TRANSMEMBRANE PROTEINS

Another approach to the structure and function of transmembrane proteins, particularly integrins, is the use of chimeric transmembrane proteins in which the cytoplasmic region from one transmembrane protein is covalently fused to the extracellular and transmembrane region of another transmembrane protein. Unlike the protein constructs discussed above, this fusion is accomplished by genetic engineering techniques, and each

of the two subunits of the heterodimeric chimeric protein is synthesized by a normal protein synthesis mechanism in which the chain is synthesized as one unit from the amino terminus to the carboxy terminus.

In these chimeric transmembrane proteins, the cytoplasmic domain of an integrin can be fused to the transmembrane and extracellular domains of a non-integrin transmembrane protein, such as the Tac subunit of the human IL-2 receptor. Alternatively, the cytoplasmic domain of one integrin can be fused to the transmembrane and extracellular domains of another integrin.

Examples include a chimeric integrin protein comprising the extracellular and transmembrane domains of the Tac subunit of the IL-2 receptor covalently linked to the cytoplasmic domain of integrin β_3, a chimeric integrin comprising the transmembrane and extracellular domains of the Tac subunit of the IL-2 receptor and the cytoplasmic domain of integrin α_{IIb}.

In the chimeric integrin in which the cytoplasmic domain is that of integrin β_3, the amino acid in the cytoplasmic domain of integrin β_3 that is amino acid 752 in the complete integrin β_3 molecule can be mutated from a serine residue to a proline residue. This mutation is associated with defective inside-out signalling in intact integrins.

Another aspect of the present invention is a heterodimeric chimeric integrin in which the extracellular and transmembrane domains of human integrin $\alpha_{IIb}\beta_3$ are joined to the cytoplasmic domains of human integrin $\alpha_5\beta_1$, so that the α subunit has the extracellular and transmembrane domains of α_{IIb} and the cytoplasmic domain of α_5, and the β subunit has the extracellular and transmembrane domains of β_3 and the cytoplasmic domain of β_1.

Chimeric integrins can be constructed in a number of vectors well known in the art, such as the CDM8 vector (A. Aruffo & B. Seed, Proc. Natl. Acad. Sci. USA 84:8573-

8577 (1987)), or can be constructed in the CMV-IL2R vector for production of chimeric integrins in which the extracellular and transmembrane domains are derived from Tac (S.E. LaFlamme et al., J. Cell Biol. 117:437-447 (1992)). Alternatively, other vectors known in the art can be used.

General cloning techniques are well known in the art and need not be described further here; they are described, for example, in J. Sambrook et al., "Molecular Cloning: A Laboratory Manual" (Cold Harbor Laboratory Press, Cold Spring Harbor, New York, 1989), vol. 1-3; B. Perbal, "A Practical Guide to Molecular Cloning" (John Wiley & Sons, New York, 1988); D.V. Goeddel, ed., "Gene Expression Technology" (Methods in Enzymology, vol. 185, Academic Press, Inc., San Diego, 1991) and other sources. In general, methods for selecting the vector, choosing the appropriate cleavage points for restriction endonucleases, ligating the appropriately cleaved or synthesized segment into the vector, introducing the chimeric vector containing the DNA segment to be expressed into a suitable host cell by techniques such as transfection, lipofection, and other techniques, and expressing the desired protein are well known in the art and described in these references. Therefore, they need not be elaborated upon further here.

Mutations can be introduced into the chimeric integrins by site-specific mutagenic techniques, such as those involving the polymerase chain reaction technique (J.C. Loftus et al., Science 249:915-918 (1990)), as well as other site-specific mutagenic techniques described in the cloning references given above.

The present invention also encompasses nucleic acid sequences encoding chimeric integrins as described above, and nucleic acid sequences operatively linked to at least one control element for transcription of the nucleic acid sequence. Suitable control elements are well known in the art and include promoters and enhancers. These are described in the references cited above, namely Maniotis

et al., Perbal, and Goeddel. Promoters include the Rous sarcoma virus LTR promoter, the SV40 early promoter, the herpes simplex thymidine kinase promoter, the mouse mammary tumor virus LTR promoter, the adenovirus major late promoter, the murine metallothionein promoter, the avian sarcoma virus promoter, the human cytomegalovirus immediate early promoter, the interferon β promoter, the Drosophila heat shock protein 70 promoter and the hMTII promoter. Other promoters and enhancer elements are well known in the art and need not be described further here.

A further aspect of the invention is vectors comprising the nucleic acid sequence coding for a chimeric integrin operatively linked to at least one control element. Vectors according to the present invention are capable of transfecting at least one eukaryotic host for expression of the chimeric integrin encoded by the nucleic acid sequence. Vectors are well known in the art and need not be described further here; they are described in the references cited above. Further examples of vectors are given in the examples.

These chimeric integrins are capable of blocking activation of human integrins. A method for blocking the activation of a human integrin can comprise the step of expressing a chimeric integrin whose extracellular and transmembrane domains are derived from the Tac subunit of the human IL-2 receptor and whose cytoplasmic domain is the cytoplasmic domain of the human β_1 integrin in a quantity sufficient to inhibit high affinity ligand binding by the cellular integrin. This is a competitive method. The preparation of the chimeric integrin whose extracellular and transmembrane domains are derived from the Tac subunit and whose cytoplasmic domain is the cytoplasmic domain of the human β_1 integrin is described in Example 4.

As another aspect of the present invention, a similar method can be used for blocking the activation of a human integrin comprising the step of expressing a chimeric integrin comprising the extracellular

transmembrane domains of the Tac subunit covalently linked to the cytoplasmic domain of integrin β_3.

These techniques block inside-out signal transduction and represents an alternative way of modulating integrin activity other than using small molecular weight competitive inhibitors of ligand binding to the extracellular domain (M.H. Ginsberg et al., J. Biol. Chem. 260:3931-3936 (1985); M.J. Humphries et al. (1986), *supra*). The method described herein involves inhibition of binding to the intracellular domain. Since integrin activation involves cell-type specific factors, such inhibition could be cell type-specific. This provides a totally new method for controlling integrin behavior in disease states such as thrombosis, inflammation, and tumor invasion and metastasis.

The present invention is illustrated by the following Examples. The Examples are for illustrative purposes only and are not intended to limit the invention.

EXAMPLES

Example 1

Synthesis of Type I Protein Construct and Other Model Peptides

The Type I protein construct shown schematically in Figure 1B and in Figure 2 was synthesized, along with the following molecules used as controls: (1) a molecule that consists of the cytoplasmic domain of integrin α_{IIb} with the carboxyl terminal amino acid changed from glutamic acid to glutamine to reflect likely *in vivo* amidation and the cytoplasmic domain of integrin β_3 covalently joined through a thioether linkage, but without the helical segments (MP-2 in Figure 2); (2) a molecule consisting of a helix dimer covalently joined through a thioether linkage without the cytoplasmic domains (Helix-Dimer in Figure 2). The Type I protein construct is identified in Figure 2 and in other figures as MP-1 as an alternative designation.

Also, the cytoplasmic tails of α_{IIb} and β_3 were

synthesized separately for conformation studies (see Example 2).

Materials. Boc-amino acids ($_t$-butyloxycarbonyl-amino acids) and 1-hydroxybenzotriazole tetramethyluronium hexafluorophosphate (HBTU) were obtained from Novabiochem (San Diego, CA). Pre-loaded Boc-amino acid (4-carboxamidomethyl)-benzylester-copoly (styrene-divinylbenzene) resins (Boc-amino acid-OCH_2-Pam-resins) were obtained from Applied Biosystems (Foster City, CA). The resin 4-methylbenzhydrylamine-resin (Lot No. 023863) was from Peninsula Labs (San Carlos, CA).

Other reagents used included N,N-dimethylformamide, and HPLC-grade acetonitrile. All other reagents were AR grade.

Peptide Synthesis. All peptides were synthesized manually according to the *in situ* neutralization/HBTU activation protocol for Boc solid phase chemistry as described previously (M. Schnölzer et al. (1992), *supra*). Coupling yields were monitored by the quantitative ninhydrin determination of residual free amine (V.K. Sarin et al., Anal. Biochem. 117:147-157 (1981)). Peptide-α-carboxamides were constructed on a 4-methylbenzhydrylamine-resin, and all other peptides synthesized on appropriate Boc-aminoacyl-OCH_2-Pam-resins. Where required, the bromoacetyl group was introduced at the N_α-terminal of a peptide by coupling as the preformed symmetric anhydride. In all cases, side-chain protecting groups were removed and the peptides cleaved from the resin by treatment with liquid HF containing 4% p-anisole for one hour at 0°C. Crude peptide products were precipitated and washed with dimethyl ether before being dissolved in aqueous acetic acid (10-30%) and lyophilized.

Reverse Phase HPLC. Analytical and semipreparative gradient HPLC were performed on a Rainin dual pump high pressure mixing system with 214 nm detection. Semipreparative HPLC was run on a Vydac C_{18} column (10 micron, 10 x 250 mm) at a flow rate of 3 ml/min.

Analytical HPLC was performed on a Whatman C_{18} column (5 micron, 4.0 x 140 mm) at a flow rate of 1 ml/min. Preparative HPLC was performed on Waters Prep 4000 system fitted with a Waters 486 tunable absorbance detector. Preparative HPLC was run on a Vydac C_{18} column (15-20 micron, 50 x 250 mm) at a flow rate of 30 ml/min. All runs used linear gradients of 90% acetonitrile plus 0.1% TFA versus 0.1% aqueous TFA.

Peptide Purification and Characterization. Crude peptides were dissolved in aqueous acetonitrile containing 0.1% TFA and purified by either semipreparative or preparative HPLC. Peptides containing cysteine residues were dissolved in HPLC buffers containing 5 mM dithiothreitol (Cleland's Reagent). All purified peptides were characterized by ion-spray mass spectrometry (see Table 1).

Mass Spectrometry. Ion-spray mass spectrometry of crude and purified peptide segments was performed on a API-III quadrupole ion-spray instrument (Sciex, Toronto) as described (M. Schnölzer et al. (1992), *supra*). Expected masses of peptide targets were determined using MacProMass.

Synthesis of Model Protein MP-1 (Protein Construct Type I). Ligation was initiated by combining Cys-helix-α_{IIb} (6.9 mg, 1.26 μmol) and BrAc-helix-β_3 (7.3 mg, 0.84 μmol) in 1.4 ml of 95% DMF/0.1 M sodium phosphate, pH 7.0 at 25°C. The ligation reaction was monitored by reverse-phase HPLC (20-50% CH_3CN over 30 minutes) and by ion-spray mass spectrometry. The ligation reaction was terminated after 12 hours and the product purified by preparative HPLC using a gradient of 30-60% CH_3CN over 60 minutes. The lyophilized purified product was characterized by ion-spray (calcd 14189.3 (monoisotopic), 14198.9 (average isotope composition), found 14194 ± 1.8)) (Table 1). The synthesis of model protein MP-1 is depicted schematically in Figure 3.

TABLE 1

MASS SPECTROMETRIC CHARACTERISTICS OF SYNTHETIC PEPTIDES

Peptide	Calculated (monoisotopic)	Calculated (av. isotope composition)	Found
αIIb	2434.2	2435.6	2434.0 ± 0.5
Cys-αIIb	2537.2	2538.7	2537.0 ± 0.9
Helix-αIIb	5452.8	5456.1	5454.3 ± 1.1
Cys-Helix-αIIb	5555.8	5559.2	5558.2 ± 1.4
β_3	5572.5	5577.2	5576.1 ± 0.9
BrAc-β_3	5693.8	5698.2	5697.3 ± 1.4
BrAc-Helix-β_3	8730.1	8736.6	*8715.4 ± 1.6
BrAc-Helix	3157.6	3160.4	3158.3 ± 0.6
Cys-Helix	3140.7	3142.6	3141.2 ± 0.5

* dehydration adduct

Synthesis of Model Protein MP-2. Cys-α_{IIb} (8.01 mg, 3.16 μmol) and BrAc-β_3 (11.78 mg, 2.1 μmol) were combined in 1.98 ml of 95% DMF/0.1 M sodium phosphate, pH 7.0 at 25°C. The ligation reaction was monitored by reverse-phase HPLC using a gradient of 20-50% CH_3CN over 30 minutes. Peaks were collected and analyzed by ion-spray mass spectrometry. The ligation reaction was terminated after 18 hours and the product purified by preparative HPLC using a gradient of 20-50% CH_3CN over 60 minutes. The lyophilized purified product was characterized by ion-spray mass spectrometry (calcd 8151.1 (monoisotopic), 8156.9 (average isotope composition), found 8151.4 ± 1.0)).

Synthesis of Helix-Dimer. Ligation was carried out by combining BrAc-[G.(K.L.E.A.L.E.G.)$_4$].-NH_2 (SEQ ID NO: 13) (16.2 mg, 5.3μmol) and Cys-[G.(K.L.E.A.L.E.G.)$_4$].-NH_2 (SEQ ID NO: 14) (13.5 mg, 4.3 μmol) and 1.5 ml of 95% DMF/0.1 M sodium phosphate, pH 7.0 at 25°C. The ligation reaction was monitored by reverse-phase HPLC (5 μl aliquots) using a gradient of 35-50% CH_3CN over 30 minutes. Peaks were collected and analyzed by ion-spray mass spectrometry. The ligation reaction was thus observed to be complete after 2.5 hours and the product purified by semipreparative HPLC using a gradient of 35-50% CH_3CN over 30 minutes. The lyophilized purified product was characterized by ion-spray mass spectrometry (calcd 6219.0 (monoisotopic), 6223.0 (average isotope composition), found 6219.4 ± 0.8).

In designing model protein MP-1, i.e., a Type I protein construct, for the cytoplasmic and transmembrane domains of the $\alpha_{IIb}\beta_3$ receptor, various structural constraints were considered. The cytoplasmic tails of the $\alpha_{IIb}\beta_3$ receptor emerged from the plasma membrane at defined points resulting in the shorter α_{IIb} tail being staggered relative to the longer β_3 tail (see Figure 1). Moreover, there is evidence to suggest that the α_{IIb} and β_3 cytoplasmic tails interact and that this interaction is

critical for proper function (J. Ylanne et al., J. Cell Biol. 122:223-233 (1993)). Such a quaternary interaction suggests that the cytoplasmic tails are located close to one another in the integrin heterodimer (Figure 1). Finally, there is indirect evidence to suggest that the transmembrane helices in $\alpha_{IIb}\beta_3$ interact within the membrane (P. Frachet et al., Biochemistry 31:2408-2415 (1992); J.S. Bennet et al., J. Biol. Chem. 268:3580-3585 (1993)) and it is possible that the structure of the cytoplasmic domain is affected by the conformation of the membrane spanning domain of $\alpha_{IIb}\beta_3$. This example demonstrates the rational design of a Type I protein construct (MP-1) and its synthesis.

The MP-1 model protein contained helical structures designed to mimic the distinct membrane-spanning domains of the two polypeptide chains in the receptor molecule, one attached to each of the cytoplasmic tails from the α_{IIb} and β_3 subunits. The sequence of the α_{IIb} and β_3 tails correspond to residue 989-1008 (M. Poncz et al., "Structure of the Platelet Membrane Glycoprotein IIb," J. Biol. Chem. 262:8476-8482 (1987)) and 716-762 (L.A. Fitzgerald et al., J. Biol. Chem. 262:3936-3939 (1987)), respectively (Figure 2). Note that there is evidence of post-translational side-chain amidation of the C-terminal glutamic acid in the α_{IIb} tail and therefore a glutamine residue was incorporated in this point in the synthetic sequence (J.J. Calvete et al. (1990), *supra*) in place of the glutamic acid produced by the translation of the nucleic acid sequence.

The sequence designed to mimic the membrane-spanning region assumes a helical conformation in these regions of the native molecule (J. Deisenhofer et al. (1985), *supra*). In order to imitate the effects of this membrane-spanning structure on the attached cytoplasmic tails, a coiled-coil structure was used, suitably modified to provide solubility under aqueous conditions. Each chain of the coiled-coil is composed of a 29 amino acid residue amphiphilic sequence, and is itself is made

up of four tandem repeats of a 7 residue core peptide [G.(.K.L.E.A.L.E.G)$_4$] (SEQ ID NO: 13). This particular sequence is derived from the prototypical coiled-coil protein tropomyosin (J. Sodek et al., J. Biol. Chem. 253:1129-1136 (1978)) and is known to form helical secondary structure in aqueous solution (S.Y.M. Lau et al., J. Biol. Chem. 259:13253-13261 (1984)). Furthermore, synthetic peptides containing polymeric assemblies of the 7 residue core sequence are known to adopt coiled-coil tertiary and quaternary structures in aqueous systems (N.E. Zhou et al., J. Biol. Chem. 267:2664-2670 (1992)).

The stability of these coiled-coils is largely the result of strong interchain hydrophobic interactions between leucine residues in the 7 residue repeat (S.Y.M. Lau et al. (1984)). By incorporating two of these amphiphilic segments into MP-1, the absolute requirement for helicity is met, but not at the price of insolubility as might be the case if the natural hydrophobic membrane-spanning sequence were used. Moreover, the tendency of these helical elements to associate to form coiled-coils may better mimic the proximity of transmembrane helices in the natural system and also ensure that a defined topology is maintained between α and β cytoplasmic tails. In other words, the coiled-coil acts as a structural template onto which the cytoplasmic domain of the integrin can be attached. This ensures that the two cytoplasmic tails are staggered with respect to one another in a manner that approximates the intact protein (see Fig. 1). Helical coiled-coil structures have previously been used as templates for the presentation of a small peptide motif (M. Engel et al., Biochemistry 30:3161-3168 (1991)).

To better gauge the structural impact of a coiled-coil template, a sequence of reference compounds was also designed. This included a second model protein, designated MP-2, containing the cytoplasmic tails of the $\alpha_{IIb}\beta_3$ receptor linked together in a head-to-head manner

with no coiled-coil structure template. The individual cytoplasmic tails of the two subunits, on their own and with the additional amphiphilic sequence on their amino-termini were also deemed useful control molecules, as were the head-to-head linked coiled-coil segments themselves. The primary structures of all of the component peptides and the resulting assemblies are shown in Figure 2.

The model proteins illustrated in Figures 1 and 2 have a somewhat unusual architecture, namely two carboxyl-termini. This feature renders them not directly accessible via either recombinant technology or via standard chemical polypeptide synthesis. A novel strategy must be employed, such as chemoselective ligation (M. Schnölzer & S.B.H. Kent, Science 256:251-225 (1992)). This strategy involves the chemical dovetailing of two fully unprotected peptide segments by a chemoselective ligation reaction to form the mature target compound. The selectivity of this ligation reaction is imposed by incorporating unique, mutually reactive groups, one within each peptide segment to be joined. In the case of MP-1, each half of the target molecule was individually constructed, and then these two intermediates were joined together through the N-termini of the helices with a thioether linkage. In this case, the ligation chemistry takes advantage of the fact that neither cytoplasmic tail of the $\alpha_{IIb}\beta_3$ receptor contains a cysteine residue. Thus, by including a unique cysteine residue with its nucleophilic sulfhydryl at the N-terminus of one half of MP-1 and an electrophilic bromoacetyl moiety at the N-terminus of the other half, the two pieces can be chemically dovetailed in the desired manner (see Fig. 3). The same principle applies to the synthesis of the control protein, MP-2, in which the two cytoplasmic tails are directly linked together with no intervening helical regions.

The feasibility of this approach was tested by studying the chemical ligation of two derivatives of the

amphiphilic helical peptide. Thus, both the nucleophilic peptide segment Cys-G. $(K.L.E.A.L.E.G.)_4$ and the electrophilic component BrAc-G. $(K.L.E.A.L.E.G)_4$ were synthesized as peptide-amides and their mutual reactivity investigated. The choice of pH for the reaction is critical since it must be sufficiently high to ionize the side chain -SH of Cys and render it nucleophilic. Preliminary studies have revealed that ligations could not be brought about between cysteine-containing peptides and bromoacetylated peptides at low pH. However, too high a pH will also deprotonate the ϵ-amino groups in lysine side-chains (pKa ~ 10.5) causing them to react with electrophiles such as the bromoacetyl group, and leading to undesired reaction products. A compromise must be struck between these two effects (R. Wetzel et al., Bioconjugate Chem. 1:114-122 (1990)).

Initially, the ligation reaction was performed under aqueous conditions (pH 7.0, 0.1 M phosphate, 10 mg/ml in each reactant), and the progress of the ligation reaction was monitored by analytical HPLC (Fig. 4a). Individual HPLC peaks were collected and examined by mass spectrometry. Two chemical reactions were observed under these conditions. The desired ligation product (thioether-linked helix dimer) was formed in significant amounts after only 45 minutes. However, substantial amounts of the disulfide homodimer of the cysteinyl-component also formed under these conditions. Formation of this oxidized species effectively protects the reactive sulfhydryl group thereby substantially reducing the ligation yield. Adjustment of the reaction conditions was therefore needed to enhance the nucleophilic ligation reaction relative to the unwanted oxidation reaction. Air oxidation of a mercaptan to a disulfide follows a multistep mechanism involving sulfur radical formation (T.J. Wallace et al., J. Org. Chem. 28:1311-1314 (1963)), whereas the nucleophilic attack of thiolate ion on a primary alkyl halide, such as a bromoacetyl moiety, occurs by a S_N2 mechanism. Thus, the

reaction rate for ligation should be greatly increased relative to oxidation by the use of a dipolar aprotic solvent. This is because of two effects: enhanced reactivity of the charged nucleophile due to reduced solvation, and better solvation of the charge-separated transition state (Parker, Chem. Rev. 69:1-32 (1969)).

To test this hypothesis, the ligation reaction was repeated using a solvent system composed of 95% dimethyl formamide (DMF) and 5% 0.1 M phosphate at pH 7.0, the aqueous phosphate being added to ensure thiolate formation (M. Baca & S.B.H. Kent, Proc. Natl. Acad. Sci. 90: 11638-11642 (1993)). To further favor dimerization of the cysteinyl component, the concentration of each reactant was reduced from 10 mg/ml to 5 mg/ml. Under these conditions, the nucleophilic ligation reaction proceeded extremely quickly. Considerable amounts of product were present at the earliest time points examined, after less than one minute of reaction. After 2.5 hours, the reaction had already gone to completion, as indicated by the complete disappearance of the cysteinyl component (Fig. 4b). Importantly, absolutely no disulfide-linked dimer was observed. Thus, by use of the dipolar aprotic solvent DNF, disulfide formation can be eliminated in favor of the desired ligation reaction.

The synthesis of MP-2, by reaction of H-Cys-[α_{IIb}]-OH with BrAc-[β_3]-OH, was similarly performed in 95% DMF/ 5% phosphate, pH 7.0) using a concentration of 5 mg/ml in each reactant. As expected from the model ligation studies, the reaction proceeded virtually to completion under these conditions. The ligation product, MP-2, was subsequently purified by preparative HPLC and its covalent structure confirmed by ion-spray mass spectrometry (Fig. 5a). The model protein MP-1, a Type I protein construct, was prepared by chemical dovetailing of the peptides H-Cys-G.(K.L.E.A.L.E.G)$_4$-[α_{IIb}]-OH and BrAc-G.(K.L.E.A.L.E.G)$_4$-[β_3]-OH. Chemoselective ligation was carried out in 95% DMF/ 5% phosphate, pH 7.0 using a

concentration of 5 mg/ml in each reactant. The ligation product, MP-1, was purified to homogeneity by preparative HPLC and characterized by ion-spray mass spectrometry (Fig. 5b). Significantly, this ligation reaction was observed to proceed much slower than for the model ligation studies and for the synthesis of MP-2, both described above. After 12 hours, the yield of ligated product was approximately 60%. At this point, the reaction was terminated since the residual bromoacetylated peptide had almost all been converted to the unreactive chloroacetyl adduct (as indicated by mass spectrometry), by chloride ions of unknown origin. Despite this side reaction, the formation of the Type I protein construct MP-1 proceeded at a useable rate in a substantial yield.

Example 2

Circular Dichroism and Size Exclusion Chromatography on Molecules of Example 1

Ultraviolet Circular Dichroism Spectroscopy. Far ultraviolet CD spectra were recorded on an AVIV 60DS spectropolarimeter linked to an AT & T computer. All peptide and protein samples were dissolved in 50 mM boric acid at pH 7.0 and their concentrations determined by quantitative amino acid analysis. CD spectra are presented as a plot of mean molar ellipticity per residue ($[\theta]$, deg cm^2 $dmol^{-1}$) versus wave length in 0.5 nm increments. The digitized data was plotted using the Cricket graph program on a Macintosh IIsi computer.

Calculation of Protein Helicity. The percentage of helical secondary structure within a sample was estimated using equation 1. The maximal theoretical ellipticity at 222 nm $[\theta]_{max}$ was determined using equation 2 where n is the number of residues per chain (Y.H. Chen et al., Biochemistry 13:3350-3359 (1974)).

$$\% \text{ helix} = [\theta]x_{222}/[\theta_{max}] \times 100 \quad (1)$$

$$[\theta]_{max} = -39500[1-(2.57/n)] \quad (2).$$

Size Exclusion Chromatography. Chromatography was performed on a Pharmacia FPLC System using a Superdex

column (240 mm x 12 mm (i.d.)). Peptide and protein samples (100 μl of 50 μM solution) were eluted at a flow rate of 0.5 ml/min with 0.1 M phosphate buffer at pH 7.0 containing 0.5 M NaCl. Sample elution was monitored either by absorbance at 214 nm or by fluorescence at 235 nm with excitation at 280 nm. A series of peptides of varying lengths derived from interleukin-8 and the fibronectin tenth type III module, as well as the α_{IIb} and β_3 tails were used as molecular weight standards.

The far UV spectra of the cytoplasmic tails of both α_{IIb} and β_3 were taken in boric acid, pH 7.0 at 25°C (Figure 6a). Significantly, in each case, the absence of distinct minima at 208 and 220 nm in the CD spectra revealed that neither cytoplasmic tail contained appreciable helical secondary structure in aqueous solution. However, in other respects, the two CD spectra were strikingly different. The spectrum of the 47 amino acid β_3 cytoplasmic tail contained a strong minimum at 203-205 nm as opposed to that of the 20 amino acid α_{IIb} cytoplasmic tail which was characterized by a small maximum at 219-220 nm and a weak negative ellipticity at 235 nm (Fig. 6a).

The far UV CD of the α_{IIb} cytoplasmic tail clearly indicates a non-helical structure. The spectrum has many features in common with classical random-coil polyamino acids (R. Townend et al., Biochem. Biophys. Res. Comm. 23:163-169 (1966)). The primary sequence of the α_{IIb} tail contains a series of six acetic amino acids in a row (Glu-Glu-Asp-Asp-Glu-Glu), and at pH 7.0 there is potential for electrostatic repulsion between neighboring carboxylates, perhaps accounting for this disordered structure. The β_3 cytoplasmic tail exhibits a far UV CD spectrum which differs significantly from the classical random-coil spectrum exhibited by the α_{IIb} cytoplasmic tail. The spectrum indicates that the β_3 cytoplasmic tail also contains little if any α-helix. Thus, while both isolated cytoplasmic tails are clearly non-helical, these

spectral differences point to the cytoplasmic tails having dissimilar structural properties.

The far UV spectra of both H-G.(K.L.E.A.L.E.G)$_4$-[α_{IIb}]-OH (helix-α_{IIb}) and H-G.(K.L.E.A.L.E.G)$_4$-[β_3]-OH (helix-β_3) were markedly different from the corresponding cytoplasmic domains alone (Fig. 6b). Both helix-α_{IIb} and helix-β_3 exhibited bimodal UV CD spectra with minima at 208 and 222 nm, indicative of helical secondary structure (G. Holzwarth & P. Doty, J. Am. Chem. Soc. 87:218-228 (1965)). The protein helicity estimated from [θ_{222}] using equations 1 and 2 was 44% for helix-α_{IIb} and 23% for helix-β_3 (Table 2). The absence of detectable helicity in the individual α_{IIb} and β_3 cytoplasmic tails suggest that the α-helix is restricted to the N-terminal pro-(coiled-coil) sequence in helix-α_{IIb} and helix-β_3. The ratio of $[\theta]_{222}/[\theta]_{208}$ has previously been used to assess the number of helical strands within a molecule (S.Y.M. Lau et al. (1984), *supra*). A value for $[\theta]_{222}/[\theta]_{208}$ of around 0.8 is associated with single-stranded α-helix, whereas a value of around unity is suggestive of a two-stranded coiled-coil. By this measure, helix-α_{IIb} with a $[\theta]_{222}/[\theta]_{208}$ value of 1.02 would appear to contain a coiled-coil structural unit. Similarly, helix-β_3 ($[\theta]_{222}/[\theta]_{208}$=1.20) would also contain a coiled-coil by this criterion.

TABLE 2

PERCENT HELICITY OF SYNTHETIC PROTEINS CALCULATED FROM FAR UV CD DATA USING EQUATIONS (1) AND (2)

Sample	Protein Helicity	$\theta_{222}/\theta_{208}$
Helix-αIIb	44%	1.02
Helix-β_3	23%	1.25
Helix-Dimer	85%	1.00
MP-1	74%	1.08

The moderate protein helicity exhibited by the helix-[cytoplasmic tail] molecules would seem to preclude the intramolecular formation of a coiled-coil architecture, implying that the molecules exist as homodimeric coiled-coils. To test this hypothesis, size exclusion chromatography was performed on both molecules. When compared with a number of peptide standards of varying length, the retention volumes of helix-α_{IIb} (monomer = 5.4 kDa) and helix-β_3 (monomer = 8.7 kDa) corresponded to apparent masses of approximately 9.9 and 16 kDa respectively. These data support the existence of homodimeric species. The putative intermolecular coiled-coils within these homodimers presumably involved a noncovalent association of the two N-terminal amphiphilic helices stabilized by hydrophobic interactions.

The CD spectrum of the 60 amino acid helix-dimer (Fig. 2) was indicative of an α-helical coiled-coil protein (Fig. 7). Both the estimated protein helicity of 85% (Table 2) and the $[\theta]_{222}/[\theta]_{208}$ value of exactly 1.00 confirmed the presence of this tertiary structure in the covalently linked molecule. This observation is of key importance since the presence of a coiled-coil in such a parallel architecture is a fundamental design feature of the model protein MP-1, a Type I protein construct. The far UV CD spectrum of MP-2, which lacks the amphiphilic helical regions, reveals that the 69 residue molecule does not contain substantial helical secondary structure (Fig. 7), an observation which is entirely consistent with the CD studies on the individual cytoplasmic tails.

The far UV CD spectrum of the 126 residue model protein MP-1, a Type I protein construct, contained minima at 208 and 220 nm, again indicating the presence of helical secondary structure (Fig. 7).

The calculated $[\theta]_{222}/[\theta]_{208}$ value of 1.08 suggests that as expected the molecule does contain an area of α-helical coiled-coil. The retention volume of the MP-1 (monomer = 14 kDa) molecule on gel permeation chromatography corresponded to an apparent mass of 18

kDa. This suggests that the molecule exists as a monomer in aqueous solution, implying that this coiled-coil is intramolecular. In stark contrast to the non-helical MP-2, the protein helicity of MP-1 was estimated to be some 74%. Based on the covalent structure of the MP-1 molecule, the expected coiled-coil region makes up only 47% of the amino acid content of the molecule. Thus, the 74% observed helicity (Table 2) indicates that helicity extends beyond the expected amphiphilic coiled-coil region of the molecule into the cytoplasmic domains. It is significant that this additional helicity was not observed for either helix-α_{IIb} or helix-β_3, both of which appear to exist as homodimers containing coiled-coils. This argues that the coiled-coil unit is not itself sufficient to induce helicity into the cytoplasmic tails. Thus, the additional helicity which is unique to the MP-1 molecule must be a result of the α_{IIb} and β_3 tails being present in a staggered and parallel orientation within this molecule.

Example 3

Fluorescence Quenching Studies on Molecules Synthesized in Example 1

For fluorescence quenching studies, protein fluorescence was measured on a Jasco FP-777 spectrofluorometer. The quenching buffer comprised 50 mM NH_4OAc, pH 6.0 containing 0.16 M KCl and 0.1 mM $Na_2S_2O_3$. The KI stock solution (1.5 M) was prepared gravimetrically using quenching buffer. Protein stock solutions (typically 100 μM) were similarly prepared with quenching buffer. For each data point in the quenching experiment, 200 μl of protein stock was added to the appropriate amount of KI stock and the final volume made to 1 ml with quenching buffer. All solutions were incubated for one hour at 25°C before measurements were taken. Fluorescence emission was monitored at 25°C at 350 nm with excitation at 278 nm (slit width for both excitation and emission = 10 nm) and expressed as F^0/F where F^0 and F are the fluorescence of the protein in the

absence and presence of quencher respectively. The data are presented as direct Stern-Volmer plots of F^0/F versus quencher concentration. Stern-Volmer constants K_Q were calculated using equation 3:

$$F^0/F = 1+K_Q[I^-] \quad (3).$$

Quenching of protein fluorescence by addition of external heavy atoms such as the iodide ion can provide valuable information regarding the local structural environment of the tryptophan or tyrosine side chain fluorophores (S.S. Lehrer & P.C. Leavis, Methods Enzymol. 49:222-236 (1978)). Quenching occurs as a result of diffusion controlled encounters of a heavy atom with the fluorophore, and so can be related to the solvent exposure of the residue. The MP-1 model protein contains a single tryptophan fluorophore located 24 residues from the C-terminal of the β_3 tail. This can thus be used as a convenient structural probe into this area of the protein.

Iodide quenching data was obtained for MP-1, MP-2, β_3 cytoplasmic tail, and the β_3 tail mixed with a stoichiometric amount of the α_{IIb} cytoplasmic tail. The data sets are presented as plots of F^0/F versus iodide concentration, where F^0 and F are fluorescence in the absence and presence of quencher respectively (Fig. 8). In each case, the fluorescence quenching appears to follow the simple Stern-Volmer relationship (equation 3) suggesting that the quenching of the tryptophan fluorescence is not complicated by energy transfer to tyrosine residues, the nearest of which is 8 residues downstream. The resulting Stern-Volmer constants K_Q, calculated from the graphs using equation 3, reveal a significant difference between MP-1 and the other molecules studied (Table 3).

TABLE 3

FLUORESCENCE QUENCHING STUDIES OF MP-1 AND RELATED CONTROL COMPOUNDS

Sample	K_Q (M^{-1})
β_3	4.3 ± 0.2
$\beta_3 + \alpha IIb$	4.7 ± 0.3
MP-2	4.4 ± 0.2
MP-1	0.84 ± 0.06

There appears to be little difference in the solvent exposure of the lone tryptophan of MP-2 and the β_3 cytoplasmic tail alone as indicated by the similar values of K^Q calculated for each. Furthermore, addition of the α_{IIb} cytoplasmic tail to a solution of the β_3 cytoplasmic tail has no apparent effect on the tryptophan quenching. The MP-1 fluorescence, on the other hand, is quenched to a considerably lesser extent, indicating that the single tryptophan in this Type I protein construct is substantially protected from solvent. In light of the measurements on control compounds, it appears that the solvent shielding can be attributed to tertiary interactions between α_{IIb} and β_3 cytoplasmic tails with the formation of a hydrophobic core involving the tryptophan. This suggests that the Type I protein construct, MP-1, has a defined secondary structure.

Example 4

Construction and Expression of Isolated Integrin Cytoplasmic Domains and Chimeric Molecules

cDNA Constructs. The construction of the chimeric integrins in the CDM8 vector (A. Aruffo & B. Seed (1987), *supra*, has been described (T.E. O'Toole et al., Blood 74:14-18 (1989); T.E. O'Toole et al., Science 254:845-847 (1991); T.E. O'Toole et al., J. Cell Biol. 124: 1047-1059 (1994)).

The Tacβ^1 and α_5 chimeras in the CMV-IL2R vector have been described (S.E. LaFlamme et al., "Regulation of Fibronectin Receptor Distribution," J. Cell Biol. 117:437-447 (1992)). Additional Tac chimeras were prepared by amplifying the β_3(S752P) cytoplasmic domain in the CD3a(S^{752})P vector (T.E. O'Toole et al. (1994), *supra*) using primers 5'-G-G-A-A-G-C-T-T-C-T-C-A-T-C-A-C-C-A-T-C-C-A-C-G-A-C-C-3' (SEQ ID NO: 6) and 5'-G-C-C-T-C-G-A-G-T-T-A-A-G-T-G-C-C-C-C-G-G-T-A-C-G-T-G-A-3' (SEQ ID NO: 7) by use of the polymerase chain reaction as described (J.C. Loftus et al., Science 249:915-918 (1990)). The α_{IIb} cytoplasmic domain was amplified from CD2b (T.E. O'Toole

et al. (1989), *supra*) using the primers 5'-G-G-A-A-G-C-T-T-G-G-C-T-T-C-T-T-C-A-A-G-C-G-G-A-A-C-3'(SEQ ID NO: 8) and 5'-C-C-C-T-C-G-A-G-C-T-T-G-G-A-G-G-C-A-A-C-T-T-G-T-T-G-G-3' (SEQ ID NO: 9). The isolated products were ligated into the CMV-IL2R vector between Hind III and Xho I sites.

Cell Culture and Transfection. Chinese hamster ovary (CHO) cells were transiently transfected by lipofection. CHO cells were grown in modified minimal essential medium DMEM (BioWhittaker, Walkersville, MD) containing 10% fetal bovine serum, 1% non-essential amino acids, 2 mM L-glutamine, 100 units/ml penicillin, and 100 μg/ml streptomycin. One to three days after the last passage, the cells were washed once with DMEM. 2 μg each of integrin α and β cDNA were mixed with 5 μg Tac-β_3 (Panels B and C) or Bluescript® KS (Stratagene, San Diego, CA) plasmid DNA (Panel A) and 20μl of a 3:1 liposome emulsion formed from 2,3-dioleyloxy-N-[2-({2,5-bis[(3-aminopropyl)amino]-1-oxypentyl}amino)ethyl]-N,N-dimethyl-2,3-bis(octadecanyloxy)-1-propanaminium trifluoroacetate and dioleyl phosphatidylethanolamine (Lipofectamine®, BRL, Bethesda, MD) and made up to a final volume of 200 μl with DMEM. After 10 min at room temperature, the mixture was added to the cells in a 100 mm tissue culture plate followed by addition of 3.8 ml DMEM. The cells were returned to the incubator for 6 h. After this period, the cells were washed once with complete medium. The cells were then grown in complete medium that was changed after 24 hours. COS 7 cells were transfected by a similar protocol. In some experiments, CHO and COS 7 cells were transfected by electroporation as described (T.E. O'Toole et al. (1994), *supra*). Cells were routinely analyzed 48 h after transfection.

Flow Cytometric Analysis of Integrin Affinity. PAC1 binding was analyzed by two color flow cytometry. Single cell suspensions were obtained by harvesting with 3.5 mM EDTA, incubating for 5 min in 0.1 mg/ml TPCK trypsin (Worthington) and diluting with an equal volume of

Tyrode's solution (M.H. Ginsberg et al., Blood 55: 661-668 (1980)) containing 10% fetal calf serum and 0.1% soybean trypsin inhibitor. After washing, 1-2 x 10^6 cells were incubated in a final volume of 50 μl containing 0.2% PAC1 ascites in the presence or absence of competitive inhibitor (either 2 mM peptide Gly-Arg-Gly-Asp-Ser-Pro (GRGDSP) (SEQ ID NO: 10), or 1μM peptide mimetic Ro43-5054 (L. Alig et al., J. Med. Chem. 35:4393-4407 (1992)). After a 30 min incubation at 22°C, cells were washed with cold Tyrode's solution and then incubated on ice with biotinylated anti-$\alpha_{IIb}\beta_3$, D57. After 30 min., cells were washed and then incubated on ice with Tyrode's solution containing 10% FITC-conjugated goat anti-mouse IgM (Tago, Burlingame, CA) and 2% phycoerythrin-streptavidin (Molecular Probes Inc., Junction City, OR). After 30 min, cells were diluted to 0.5 ml with Tyrode's solution and analyzed on a FACScan (Becton Dickinson) flow cytometer as described (T.E. O'Toole et al., Cell Regulation 1:883-893 (1990)). PAC binding (FITC staining) was analyzed only on a gated subset of cells positive for $\alpha_{IIb}\beta_3$ expression (phycoerythrin staining), indicated by the region M1 in Figure 10D. To define affinity state, histograms depicting PAC1 staining in the absence or presence of inhibitors were superimposed. A rightward shift in the histogram in the absence of inhibitor is indicative of the presence of high affinity $\alpha_{IIb}\beta_3$. To obtain numerical estimates of integrin activation, an activation index (AI) was calculated defined as

$$100 \times (F_0-F_R)/F_R \quad (4)$$

where F_0 is the mean fluorescence intensity in the absence of inhibitor and F_R is the mean fluorescence intensity in the presence of the competitive inhibitor as described above. In each experiment, the percent inhibition was

$$100 \times (AI_0-AI)/AI_0 \quad (5)$$

where AI_0 is the activation index in the absence of co-transfected Tac chimera and AI is the activation index in

its presence.

It is suggested that integrin cytoplasmic domains bind to intracellular elements to regulate ligand binding affinity. Therefore, an intracellular excess of cytoplasmic domains could competitively alter affinity by affecting inside-out signaling. To test the idea, a constitutive high-affinity integrin comprised of the extracellular and transmembrane domains of human $\alpha_{IIb}\beta_3$ joined to the cytoplasmic domains of human $\alpha_5\beta_1$ (Fig. 9) was transiently expressed in Chinese Hamster Ovary (CHO) cells.

Figure 9 depicts the schematics of the integrin chimeras and Tac chimeras used here. In each chimera, a human integrin cytoplasmic domain is joined to the transmembrane and extracellular domain of human Tac, β_3, or β_{IIb}. Below are depicted the single letter amino acid sequences of the integrin cytoplasmic tails. β_1 was joined to β_3 at β_3 Phe$_{727}$. The S^{752}→P point mutation in β_3 is indicated and the residues deleted in the $\alpha_L\Delta$ cytoplasmic domain are denoted by dots.

Physiological signals are involved in the maintenance of the high affinity state of this integrin because it is cell type-specific, dependent upon metabolic processes, and requires distinct structural features of the cytoplasmic domain.

The affinity of the extracellular α_{IIb} β_3 reporter group was monitored by binding of a ligand-mimetic monoclonal antibody, PAC1 (S.J. Shattil et al., J. Biol. Chem. 260: 11107-11114 (1985); Y. Tomiyama, "A Molecular Model of RGD Ligands," J. Biol. Chem. 267:18085-18092 (1992)). The cells were also stained with an affinity-insensitive anti-$\alpha_{IIb}\beta_3$ antibody (D57) and two color fluorescence was used to select only those cells that expressed α_{IIb} β_3 for analysis of PAC1 staining. To add isolated cytoplasmic domains, the CHO cells were co-transfected with chimeras of the extracellular and transmembrane domains of Tac joined to various integrin

tails (Fig. 9). The Tac-β_1 and β_3 chimeras contained sufficient information for localization to sites of membrane-cytoskeleton association (S.E. LaFlamme et al. (1992), *supra*) and, when overexpressed, inhibit cell spreading, migration, and matrix assembly.

The results from FACS studies are shown in Figure 10. Depicted in Figure 10 are flow cytometry histograms in which fluorescence intensity is plotted on the abscissa and cell number on the ordinate. PAC binding in the absence (filled histogram) or presence (open histogram) of competitive inhibitor is depicted in Panels A, B, and C. Panel D depicts surface expression of A_{IIb} β_3 as reported by the binding of the D57 antibody in the absence of (filled histogram) or presence (open histogram) of co-transfected Tac-β_3 cDNA. M-1 indicates the region containing those cells that express the recombinant α_{IIb} β_3 construct bearing the cytoplasmic domains of β_1 and α_5. The extracellular domain of this integrin specifically bound PAC1 when it was expressed in CHO cells (Panel A). Co-transfection with Tac-β_3 blocked PAC1 binding (Panel B) but did not reduce α_{IIb} β_3 expression (Panel D). PAC1 binding to the Tac-β_3 transfected cells was restored by addition of 2 μM anti-LIBS 6 activating antibody (Panel C).

CHO cells containing the recombinant integrin bound PAC1 and addition of a peptide competitive inhibitor (M.H. Ginsberg et al., J. Biol. Chem. 260:3931-3936 (1985)) blocked binding, verifying the specificity of PAC1 binding (Fig. 10, Panel A). Co-transfection of Tac-β_3 or Tac-β_1 suppressed specific PAC1 binding (Fig. 10, Panel B). This effect was not due to disruption of assembly of the integrin, since assembly-dependent (T.E. O'Toole et al. (1989), *supra*) surface expression was not reduced by co-transfection with Tac-β_3 (Fig. 10, Panel D). Moreover, PAC1 binding was restored by addition of an anti-β_3 monoclonal antibody that "activates" α_{IIb} β_3 independent of intracellular signalling (A.L. Frelinger

III et al., J. Biol. Chem. 266:17106-17111 (1991)) (Fig. 10, Panel C). Tac-β chimeras reduced the affinity of the integrin extracellular domain in both CHO and COS7 cells transfected by either electroporation or lipofection. To exclude the possibility that Tac-β_3 caused secretion of a soluble modulator of integrin affinity, separate plates were transfected with a Tac-β chimera and the integrin. The cells transfected separately with the integrin and Tac-β_3 were then mixed and co-cultured for two days. The co-culture manifested the same specific PAC1 binding as cells transfected with the integrin alone. Thus, overexpression of these Tac chimeras blocked intracellular integrin activation but not activation induced by the binding of an antibody to the extracellular domain.

In order to quantify the inhibitory effects, the effects of transfection of varying doses of the Tac-β chimeras were examined. The results are shown in Figure 11.

Figure 11a shows the inhibition of inside-out signalling. CHO cells were transfected with 2 μg each of the α and β subunits of the integrin chimera that bears the α_5 β_1 cytoplasmic domains by lipofection. The indicated quantity of Tac chimeras bearing the cytoplasmic domains of (■) β_3, (•) β_1, or (▲) α_5 were transfected at the same time. After 48 hours, cells were harvested and analyzed for PAC1 binding as described in Figure 10. A numerical activation index (AI) and percent inhibition were calculated as described above to obtain quantitative estimates of integrin activation.

Depicted are the means ± SEM of three independent experiments for each Tac chimera. Both Tac-β chimeras were inhibitory while the Tac-α_5 chimera was not. Tac-α_{IIb} and Tac without a cytoplasmic domain produced results similar to those for Tac-α_5.

Figure 11B shows the expression of the Tac chimeras.

The expression of Tac chimeras in each of the transfections shown in Panel A of Figure 11 was assessed by the binding of anti-Tac (7G7B6) (S.E. LaFlamme et al. (1992), *supra*). Data are expressed as mean ± SEM of the percentage of cells positive for 7G7B6 from the same three independent determinations depicted in Panel A. Analysis of the results shown in Figures 11A and B show that the β_3 and β_1 chimeras were roughly equipotent. In contrast, a chimera of Tac with the cytoplasmic domain of α_5 lacked inhibitory activity (Fig. 11A) even though it was expressed as well as the Tac-β constructs (Fig. 11B). Similar results were obtained with Tac joined to α_{IIb} or lacking a cytoplasmic domain. When the co-transfected cells were double stained for α_{IIb} β_3, then Tac, only about 70% of the cells expressed both markers. This may account for the failure of the Tac β constructs to achieve 100% inhibition of integrin activation. Consequently, the β subunit cytoplasmic sequences were responsible for inhibitory activity.

To determine if there were distinctive sequence requirements for β subunit cytoplasmic domain of inhibition, a Tac-β_3 ($S^{752}P$) was constructed. The β_3 ($S^{752}P$) mutation is associated with defective inside-out signaling in intact integrins (Y. Chen et al., Proc. Natl. Acad. Sci. USA 89:10169-10173 (1992)).

The results with chimeras carrying this mutation are shown in Figure 12. In Figure 12A, the effects of varying quantities of (▲) Tac-β_3 ($S^{752}P$) and (■) Tac-β_3 on PAC1 binding were compared. The experimental procedures were identical to those employed in the results shown in Figure 10. In Figure 12B, expression of Tac-β_3 ($S^{752}P$) was shown by measuring binding of the antibody 7G7B6 to the same cells shown in Panel A using the experimental procedures described in the experiments shown in Figure 10B; (▲) Tac-β_3 ($S^{752}P$); (■) Tac-β_3. These results show that when overexpressed as a Tac chimera, this mutated cytoplasmic domain lacked inhibitory activity (Fig. 12A).

Nevertheless, it was expressed to at least the same extent as a wild-type β_3 chimera (Fig. 12B). Thus, a point mutation reduces the capacity of the β_3 cytoplasmic domain to block activation. Consequently, the inhibitory activity of the β_3 cytoplasmic domain is structurally specific.

Mutations that delete portions of the conserved, membrane-proximal Gly-Phe-Phe-Lys-Arg (GFFKR) (SEQ ID NO: 5) sequence in the α subunit result in high affinity ligand binding to α_{IIb} β_3 that is independent of cellular metabolism, cell type, and the β cytoplasmic domain. Thus, activation of mutants of this type, such as the one containing the $\alpha_L\Delta$ cytoplasmic domain (Fig. 9) does not require the cellular signalling mechanism used by intact integrins. These variants have been termed "hinge mutants" to emphasize that high affinity ligand binding is their default state. To further test the idea that the β cytoplasmic domains block activation by interfering with physiological signalling mechanisms, Tac-β_1 was co-transfected with a α_{IIb} β_3 chimera bearing the $\alpha_L\Delta$ cytoplasmic domain (Fig. 9).

These results are shown in Figure 13A and 13B. The results for inhibition is shown in Figure 13A. CHO cells were transfected with 2 μg cDNA encoding β_3 joined to the β_1 cytoplasmic domain and 2 μg cDNA encoding α_{IIb} joined to the (•) α_5 or (■) $\alpha_L\Delta$ tail. The effect of co-transfection with varying doses of Tac-β_1 cDNA on PAC1 binding was assessed as described for the experiments whose results are shown in Figure 11A. For expression, in Figure 13B, the expression of Tac-β_1 when co-transfected with the (•) α_5 or (■) $\alpha_L\Delta$ chimeras described in Panel A was measured exactly as described in the experiments whose results are shown in Figure 11B.

Tac-β_1 had no effect on PAC1 binding to this integrin (Fig. 13A) even though it was expressed in the co-transfected cells (Fig. 13B). Thus, the β cytoplasmic domains inhibit activation that depends on "physiological

cellular machinery but do not reduce the ligand binding affinity of a "hinge" mutant.

There are several major implications of the results in this example: (1) The β_1 and β_3 cytoplasmic domains behave like structurally-specific competitive inhibitors of integrin activation. This result implies that there are limiting quantities of intracellular factors that bind to integrin cytoplasmic domains and modulate ligand binding affinity. (2) Tac-β_1 and β_3 act as dominant inhibitors of integrin activation. It is therefore possible to disrupt high affinity ligand binding by overexpression of these chimeras. Previous work identified extracellular integrin mutations that block ligand binding but not heterodimer assembly (J.C. Loftus et al., Science 249:915-918 (1990); Y. Takada et al., J. Cell Biol. 119:913-921 (1992)) and are dominant inhibitors of ligand binding. It will be of interest to see if different biological consequences stem from intracellular disruption of integrin activation versus blockade of the extracellular ligand binding site (R.J. Faull et al., J. Cell. Biol. 121:155-162 (1993). (3) Integrins mediate pathological processes including inflammation, tumor invasion and metastasis, and thrombosis (S.M. Albelda & C.A. Buck (1990), *supra;* M.E. Hemler (1990), *supra;* E. Ruoslahti (1991), *supra*). Small competitive inhibitors of ligand binding to the extracellular domain block cell adhesive events important in these processes (M.H. Ginsberg et al., J. Biol. Chem. 260:3931-3936 (1985); M.J. Humphries et al. (1986), *supra*). Structurally specific inhibition of integrin activation by β subunit cytoplasmic domains implies the feasibility of a novel class of intracellular inhibitors of integrin function. Since integrin activation involves cell type-specific factors, such inhibitors could be cell type-specific.

ADVANTAGES OF THE PRESENT INVENTION

The present invention provides compositions and methods for studying and controlling the structure and

activity of transmembrane proteins, particularly integrins.

Protein constructs according to the present invention have a number of applications based on the ability to maintain the cytoplasmic tails of the construct in a configuration that is equivalent or similar to the configuration predominating *in vivo* while maintaining solubility and stability in an aqueous system, namely in staggered, parallel, and proximal topology. For example, these protein constructs can be used to detect intracellular molecules capable of binding to integrins and modulating signals by inside-out signaling. Alternatively, these molecules can be used *in vivo* to disrupt or modulate inside-out signaling by binding to the cells in a manner such that the cytoplasmic domains of these protein constructs compete for intracellular molecules with the natural integrins. Because these protein constructs do not contain the extracellular ligand-binding sites of integrins, they would then disrupt inside-out signaling. This would be particularly useful in conditions in which overactivity of integrins is involved, such as inflammation, thrombosis, and malignancy. This would provide a new method of treating such conditions or their sequelae; because these molecules mimic the orientation of the natural integrins within the membrane, they would not disrupt membrane structure and would therefore be better tolerated and avoid side effects.

Additionally, protein constructs according to the present invention could be used to detect molecules capable of binding to the intracellular or cytoplasmic domain of integrins and other transmembrane molecules *in vivo*, such as by affinity chromatography.

Chimeric integrins according to the present invention can be used for blocking the activity of natural integrins *in vivo*. This activation, unlike the use of small molecule inhibitors of ligand binding by integrins, acts on the interaction between intracellular

molecules and integrins and is therefore likely to be cell type-specific. This provides yet another way of studying and modulating integrin activity *in vivo*.

Although the present invention has been described in considerable detail with regard to certain preferred versions thereof, other versions are possible. Therefore, the spirit and scope of the appended claims should not be limited to the descriptions of the preferred versions contained herein.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: The Scripps Research Institute
(B) STREET: 10666 North Torrey Pines Road
(C) CITY: La Jolla
(D) STATE: CA
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 92037
(G) TELEPHONE: (619) 554-2937
(H) TELEFAX: (619) 554-6312

(ii) TITLE OF INVENTION: STRUCTURAL MODELS FOR CYTOPLASMIC DOMAINS OF TRANSMEMBRANE RECEPTORS

(iii) NUMBER OF SEQUENCES: 20

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US 95/
(B) FILING DATE: 13-JUN-1995

(vi) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: USSN 08/260,514
(B) FILING DATE: 15-JUN-1994

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Peptide sequence recognized by integrin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Lys Gln Ala Gly Asp Val
1 5

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Ligand protein sequence recognized by integrin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Asp Gly Glu Ala
1

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Ligand sequence recognized by integrin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Glu Ile Leu Asp Val
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Ligand sequence recognized by integrin

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly Pro Arg Pro
1

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO

(v) FRAGMENT TYPE: internal